Measuring Dialect Pronunciation Differences using Levenshtein Distance

Heeringa, Wilbert Jan

Publication date:
2004

Chapter 3

Measuring segment distances discretely

A language variety allows the expression of sentences, which consist of words, which in turn consist of speech segments. When attempting to quantify distances in pronunciation between dialects, we should first make clear how the different speech segments are related to each other. In other words: the distances between the different speech segments should be determined. The relations between speech segments and the way in which distances are found are studied in this chapter and also in Chapter 4. In this chapter we focus on discrete representations of segments. These representations can be used for the corpus frequency method (see Section 2.3.2), the frequency per word method (see Section 2.3.3), and the Levenshtein distance (see Section 5.1). In our research, language varieties are mainly compared with the Levenshtein distance. Using the Levenshtein distance also allows acoustic representations of the segments to be used. The way in which distances are obtained on the basis of acoustic representations is described in Chapter 4.

In Section 3.1 we describe the different ways in which sounds can be represented. We will look at both vowels and consonants. In Section 3.2 we discuss how diphthongs are represented in the different systems. Section 3.3 discusses the way in which affricates are processed. Section 3.4 will explain how suprasegmentals and diacritics are processed. In the feature system of Hoppenbrouwers & Hoppenbrouwers rules are applied to remove redundancy in feature specifications. This is explained in Section 3.5. In Section 3.6 it is described how distances between sounds are calculated on the basis of the different representations. For each feature representation, a vowel distance matrix and a consonant distance matrix can be calculated. Once segment distances are obtained, they can be used unchanged in the Levenshtein algorithm. Another possibility is to use the logarithms of the distances. This is discussed in Section 3.7. In Section 3.8 the different feature representations are compared to each other by correlating vowel and consonant distances based on the one system with vowel and consonant distances based on the other system. Finally we draw some conclusions in Section 3.9.

When discussing the representation of segments and the processing of suprasegmentals and diacritics, we consider two data sources. One consists of a number of transcriptions of Norwegian dialects, compiled by Jørn Almberg. Each text is a translation of the fable 'The North Wind and the Sun', in Norwegian: 'Nordavinden og sola' (NOS) (see Section 7.2). The other is the *Reeks Nederlandse Dialectatlassen* (RND), a series of Dutch atlasses edited by Blancquaert and Peé (1925–1982) (see Section 9.1). For application to other dialect comparison work based on the modern IPA system, the remarks for the NOS data source should be kept in mind where there are differences between the NOS and the RND.

3.1 Representation of segments

In transcriptions words are transcribed as a series of speech segments: phones. In the simplest case the phones are not further defined. Two phones are equal or different. This simple representation is described in Section 3.1.1.

Using the phone representation it is not possible to take into account the affinity between different, but kindred segments. Methods based on phones will not regard the pair [ɪ,e] as more kindred than [ɪ,ɒ]. This problem can be solved by replacing each phonetic symbol by a bundle of features. Each feature can be regarded as a phonetic property which can be used for classifying sounds. A feature bundle is a range of feature values. Each value indicates to what extent the corresponding property is instantiated.

We present the results of experiments on three feature systems. The first feature system is described by Hoppenbrouwers and Hoppenbrouwers (2001) (H & H). This is an articulation-based system, based on Chomsky and Halle (1968). The system is interesting because the developers themselves used this system for (Dutch) dialect comparison. In Section 3.1.2 we give a more detailed description. The second feature system is developed by Vieregge et al. (1984) and Cucchiarini (1993) (V & C). This system was developed for a comparison task similar to dialectological comparison, that of checking the quality of phonetic transcriptions. This involves comparison to consensus transcriptions. This system is interesting since it is partly perception-based. We describe the system in Section 3.1.3. The last feature system is developed by Almeida and Braun (1986) (A & B), intended for checking the quality of phonetic transcriptions as well. We also included this system in our research because the well-known IPA system is directly used for finding sound distances. The system is described more detailed in Section 3.1.4.

Although each of the three feature systems seems to be a good candidate for use in dialect comparison, none of three feature systems is originally developed for the approach in which we used them. The system of Hoppenbrouwers & Hoppenbrouwers was originally meant to be used in their feature frequency method (see Section 2.3.2). Instead of comparing sounds, histograms are compared. A histogram represents for each feature the number of sounds which are positively marked for that feature in a dialect transcription. Both the system of Vieregge & Cucchiarini and the system of Almeida & Braun were developed for checking phonetic transcriptions, not for dialect comparison. Because our goal is to find dialect distances that approach human perception, the perception-based system of Vieregge & Cucchiarini may give the most promising sound distances. In Section 7.4.3 dialect distances based on the different segment representations are validated.

Note that the sounds used in the RND form a subset of the sounds of the IPA system. The RND vowels are given in Appendix A Figure A.1 and the IPA vowels are given in Figure A.2. The RND consonants are given in Figure A.3 and the IPA consonants are given in Figure A.4.

3.1.1 Phones

In transcriptions, words consist of a sequence of phones. In the IPA system each phone is noted with a basic symbol (vowels, pulmonic and non-pulmonic consonants and other symbols) and optionally supplemented with one or more suprasegmentals and/or one or more diacritics.[1] The combination of a basic symbol supplemented with some additional marks is regarded as a phone.

3.1.2 Features Hoppenbrouwers and Hoppenbrouwers

For the feature system of H & H the *Sound Pattern of English* (Chomsky and Halle, 1968) was the starting point. According to H & H this articulation-based system is considered as a standard work, generally followed in studies and modern handbooks about modern phonology (Hoppenbrouwers and Hoppenbrouwers, 2001, p. 33 and 34). The SPE system was modified and extended so that with the use of this system the RND material is done as much justice as possible. This resulted in a system with 21 features. Since we also want to dispose of a system which is suitable for (partially) processing the complete IPA system in general and the NOS data in particular, we should be able to process some suprasegmentals and diacritics which do not appear in the RND data. Therefore, it was necessary to extend the H & H system with six extra features, obtaining a total of 27 features. Both the original and the added features are given in Table 3.1. All features are initially binary, where 0 means ‘absent’ and 1 ‘present’. In the

[1] Note that in the IPA system *half-long* (ˑ) and *long* (ː) are ordered under suprasegmentals. Although this may be debatable, we will use the same ordering as starting point throughout this thesis.

Vowel features		Consonant features
vowel		consonant
front		anterior
back		coronal
round		posterior
low		laryngeal
polar		sonorant
long		voiced
peripheral		high
diphthong		continuant
nasal		lateral
		syllabic
breathy		*apical*
creaky		
toneme 1		
toneme 2		
circumflex		

Table 3.1: The features of the Hoppenbrouwers and Hoppenbrouwers system. The names of the original features were given in Dutch. In this table they are given in regular font style in English. These original features are used for the RND only. Features added for the NOS data are given in italics.

left column the vowel features are given and in the right column the consonant features can be found. The feature *nasal* is shared by both vowels and consonants.

In this section we describe the definitions of the vowels first and we give explanations about the definitions of the consonants next.

3.1.2.1 Vowels

In this section we focus on the vowels and show the relation between the IPA vowel quadrilateral and the table of H & H.

Ladefoged (1975, p. 245) uses the feature *tense* to distinguish vowels which are on the periphery of the vowel area [+tense], and the corresponding lax vowels which are slightely lower and more central [−tense]. Following Ladefoged H & H use the feature *peripheral*, which distinguishes between centralized and non-centralized vowels.

Short vowels are always specified as [−peripheral]. However, a number of vowels get [+peripheral] when they are half-long or long. From the feature table given by Hoppenbrouwers and Hoppenbrouwers (2001, pp. 37–41) it can be concluded that the [e] and [ɪ], the [ø] and [ʏ], and the [o] and [ʊ] are equal to each

		front		central		back	
polar	high	i	y			ɯ	u
	high low	e	ø			ɤ	o
polar	low	æ		a	ɶ	ɑ	ɒ

Table 3.2: IPA sounds which are (or should be) defined as peripheral sounds in the feature system of H & H. Elements left in a cell are spread, and elements right are rounded.

other if they are short. However, when they are half-long or long, the [e], [ø] and [o] are specified as [−peripheral], and the [ɪ], [ʏ] and [ʊ] as [+peripheral].

Table 3.2 presents the vowels which are or should be defined as [+peripheral] according to the feature table of H & H when they get half-long or long. [ɯ], [ɤ] and [ɶ] are not given by H & H since they are not used in the RND. The [ɒ] is not given by H & H as well, although this vowel is used in part 1 of the RND. Table 3.3 shows the vowels which are always specified as [−peripheral], regardless their length. [ɨ], [ɘ], [ɜ] and [ɐ] are not given by H & H since they are not used in the RND. The [ʏ] is in the RND and in the feature table of H & H noted as [ʌ].[2] The [ɵ] and [ɞ] are noted respectively as [ɵ̶] and [ə] in the feature table of H & H. Although the [ʉ], [ɵ] and [ɞ] are given by H & H, we did not find these vowels in the RND transcriptions we processed.

In the RND the schwa is noted as [ə], just as in the IPA system. The [ə] is defined in IPA as a half-round central vowel exactly between close-mid and open-mid. In the feature table of H & H the schwa is defined as a sound for which all features are absent, i.e., all features are set to zero. Only the features *vowel*, *sonorant*, *voiced*, *continuant* and *syllabic* are positively marked. The result is that the schwa is defined as a high central unrounded sound. So the schwa is defined as the IPA [ɘ]. Therefore, in the system of H & H the [ə] and the [ɘ] are not distinguished.

In the SPE system, intended for English segments, exactly three degrees of height can be defined using the features *high* and *low*. Defining Dutch vowels, it is necessary to be able to distinguish four degrees of height (Hoppenbrouwers and Hoppenbrouwers, 2001, p. 35). In the system of H & H this is realized with the features *low* and *polar*:

[2]In the RND the [ʌ] is introduced as a symbol representing the vowel in the Dutch word *bus* 'bus'. However, this vowel sounds approximately as the [ʏ] of the modern IPA system. For the English pronunciation of *bus* the use of the [ʌ] (as given in the modern IPA system) is correct.

		front		central		back	
polar	high			ɨ	ʉ		
	high	ɪ	ʏ	ɘ	ɵ		ʊ
	low	ɛ	œ	ɜ	ɞ	ʌ	ɔ
polar	low			ɐ			

Table 3.3: IPA sounds which are (or should be) defined as non-peripheral sounds in the feature system of H & H. Elements left in a cell are spread, and elements right are rounded.

close	-	+
close-mid	-	-
open-mid	+	-
open	+	+

The result is that the difference between close and open-mid is greater than the difference between close and open. Likewise the difference between open and close-mid is greater than between open and close. The smaller difference between the extremes (open and close) may be intended to reflect that a vowel shift is cyclic: e.g. an [ɔ] changes in a [o], an [o] changes in an [u], and next an [u] changes in an [ɒ]. We suspect that because of the 'polar' feature this feature system does not reflect the distance between segments with as much fidelity as some competitors.

Besides the features as given by Hoppenbrouwers and Hoppenbrouwers, some extra features are added. To be able to process some diacritics in the NOS data, we also needed to add the features *breathy*, *creaky*, *toneme 1*, *toneme 2* and *circumflex*.

For the use of the Levenshtein distance we will also need a definition of 'silence' (see Section 5.1). In the feature system of H & H all features can simply be defined as absent, i.e. set to 0. For vowels, this is equal to a [ə] or [ɘ] with [−vowel].

3.1.2.2 Consonants

In this section we focus on the consonants and show the relation between the IPA consonant table and the table of H & H. We will consider all pulmonic IPA consonants. For consonants treated by H & H as well as consonants not given, we show the relation with the IPA consonant table. The relation with the manner of articulation (IPA columns) and the relevant H & H features is given in Table 3.4. The relation with the place of articulation (IPA rows) and the relevant H & H features is given in Table 3.5. In the H & H system all approximants (not lateral) are defined as [+high]. For the definition of the [w] the vowel feature *round* is

IPA	ant	cor	post	lar	high
bilabial	+	-	-	-	-
labiodental	+	-	-	-	-
dental	+	+	-	-	-
alveolar	+	+	-	-	-
postalveolar	-	+	-	-	+
retroflex	-	+	-	-	+
palatal	-	+	-	-	+
velar	-	-	+	-	+
uvular	-	-	+	-	+
pharyngeal	-	-	-	-	-
glottal	-	-	-	+	-

Table 3.4: Relation between IPA columns (manner of articulation) and the relevant H & H features.

specified as [+round].[3] The distinction between voiced and voiceless sounds is defined in the same way as in the IPA table of pulmonic consonants. H & H only defined the consonants which appear in the RND.

In the IPA system the [h] represents a voiceless glottal fricative, and the [ɦ] represents its voiced counterpart. It is striking that H & H specifies the Dutch /h/ as [+voiced], just as Booij (1995) does. However, in our opinion the Dutch /h/ is voiceless. Therefore, we specify this segment as [−voiced]. This agrees with Rietveld and Van Heuven (1997, p. 395) who transcribe the /h/ in *hond* 'dog' as [h], and not as [ɦ].

Using 21 features, all RND sounds get a unique definition. However, not all sounds of the complete set of pulmonic IPA sounds are uniquely defined. So H & H write that with close to thirty features all sounds which appear in natural languages can be defined (Hoppenbrouwers and Hoppenbrouwers, 2001, p. 9).

Besides the basic features given by Hoppenbrouwers and Hoppenbrouwers, some features were added to process some diacritics. For both the RND data and the NOS data we added the feature *syllabic*. For the NOS data only we added the feature *apical*.

For the use of the Levenshtein distance we will also need a definition of 'silence' (see Section 5.1). As mentioned in Section 3.1.2.1 for 'silence' all features can simply be defined as absent, i.e. set to 0. For consonants, this is equal to the [ʔ] with [−consonant −laryngeal].

[3]In our research the [w] is regarded as a voiced bilabial approximant which can be located in the IPA table of pulmonic consonants. However, in the IPA system (revised to 1993, updated in 1996) the [w] is ordered under 'Other Symbols' and mentioned as a voiced labial-velar approximant.

IPA	nas	cons	son	cont	lat
plosive	-	+	-	-	-
nasal	+	+	+	-	-
trill	-	+	+	+	-
tap or flap	-	+	+	+	-
fricative	-	+	-	+	-
lat. fric.	-	+	-	+	+
approximant	-	-	+	+	-
lat. appr.	-	+	+	+	+

Table 3.5: Relation between IPA rows (place of articulation) and the relevant H & H features.

3.1.3 Features Vieregge and Cucchiarini

The reliability of transcriptions can be measured by determining the degree of similarity between transcriptions carried out either by the same transcriber at different times, which corresponds to the use of "reliability" in its strict sense (Bürkle, 1986), or by different transcribers, an option also left open by Vieregge et al. (1984). The validity of transcriptions is measured by comparing individual transcriptions with expert transcriptions (Vieregge et al., 1984).

In 1984 Vieregge et al. presented a feature system which was developed for checking the quality of phonetic transcriptions. This involves comparison of consensus transcriptions. The system consists of 4 multi-valued features only for vowels, and 10 multi-valued features only for consonants. Tables for vowels and consonants are given by Vieregge (1987). An advantage of Vieregge's system is that it is partly based on real measurements. The vowel system is based on experimental data which was found in the literature. The consonant system is based on a perception experiment, in which subjects were asked to give the distance between two consonants on a scale from 1 (minimal dissimilarity) to 10 (maximal dissimilarity). Next a feature system was developed, such that sound distances on the basis of features approach the perceptual distances maximally. The complete system was originally developed for Dutch.

With some extensions it may also be used for other languages as Cucchiarini (1993) showed. She extended the system so as to accommodate consonants of Limburg and Czech that had not been included earlier, as well as other sounds that probably could crop up in the transcriptions that she used. However, when expanding the system to other languages, one should be aware of the fact that different languages have different sound systems, the phonological spaces may be filled differently. Cucchiarini realizes this and writes (p. 97): "So, as it was clear that a theoretically satisfactory evaluation system was not possible, we tried to obtain a system that would at least be satisfactory from a practical point of view". The use of a Dutch Vieregge system which is extended and applied to e.g. Czech will probably reflect the perception of Dutch people listening to Czech, rather than the perception of the Czech speakers themselves.

Distance measures which are developed to assess transcriptions can also be used to quantify dialect distances. This, however, presupposes that the variable "transcriber" is kept constant by either having only one transcriber, who also undergoes reliability testing, or working with high-quality consensus transcriptions. Otherwise, there is the danger of creating so-called *Exploratorendialekte* ('explorer dialects'), i.e. "dialects" created not by differences in pronunciation but by different people transcribing them.

For our purpose of quantifying dialect distances we use a feature system which is a combination of the vowel feature system of Vieregge et al. and the consonant feature system of Cucchiarini. We first describe the vowel system and then we give a description of the consonant system.

3.1.3.1 Vowels

For the construction of the vowel feature system Vieregge et al. consulted various data in the literature. The existing literature also provided enough experimental data on the basis of which Dutch vowel distances could be found. As examples the authors refer to Nooteboom (1971), Nooteboom (1972) and Rietveld (1979). The vowel feature system of Vieregge et al. consists of four features: *advancement*, *high*, *long* and *rounded*. The possible values for the features are listed in Table 3.6. The features *advancement*, *high* and *rounded* correspond with the dimensions of the IPA vowel quadrilateral as can be seen in Table 3.7. From the values of the feature *advancement* it appears that this feature has extra weight. The authors refer to Rietveld (1979) who show that 'the proprioceptive articulatory dissimilarities can be predicted quite satisfactorily by using a traditional vowel scheme and giving extra weight to differences on the front/back dimension'. Although this statement refers to a subset of Dutch vowels, namely [i, e, ɛ, y, ø, u, o, ɔ, ɑ], Vieregge et al. assume that this finding can be applied to all Dutch vowels.

The tables of Vieregge (1987) show that Dutch [ɪ], [ʏ], [ʊ], [ə], [ɛ] and [ɑ] can only be short. The [i], [y], [u], [ɛ], [œ] and [ɔ] can only be short or long. The [e], [ø], [o] and [a] can only be half-long or long. The [ɪ] is defined as a short [e], and the [ʊ] is defined as a short [o].

For our research we extended the vowel system so that it contains all vowels of the IPA vowel quadrilateral. The result can be seen in Table 3.8. All IPA vowels are defined by analogy with the IPA vowel quadrilateral. The result is that the [ɪ] and the [e], and the [ʊ] and the [o] no longer have the same values for the features *advancement* and *high*. Now the [ʏ] is defined as a rounded [ɪ] and the [œ] as the rounded [ɛ]. The [a] is defined as a front vowel instead of a central

Feature	Value	Meaning
vowel	0	no
	1	yes
advancement	2	front
	4	central
	6	back
high	1.0	open
	1.5	near-open
	2.0	open-mid
	2.5	central
	3.0	close-mid
	3.5	near-close
	4.0	close
long	1	short
	2	half-long
	3	long
rounded	0	no
	1	yes
nasal	0.0	not nasal
	0.5	half-nasal
	1.0	nasal
breathy	0	no
	1	yes
creaky	0	no
	1	yes
toneme 1	0	no
	1	yes
toneme 2	0	no
	1	yes
circumflex	0	no
	1	yes

Table 3.6: The vowel features of Vieregge et al. and their possible values. We extended the system with some extra features, in this table given in italics. Only the first seven features in this table are used for the RND data, the last five features are added for the NOS data.

	front		central		back	
close	i	y		ʏ		u
close-mid	e/ɪ	ø	ə			ʊ/o
open-mid	ɛ			œ		ɔ
open			a		ɑ	

Table 3.7: The Dutch vowels as defined by the features of Vieregge et al.. Elements left in a cell are spread, and elements right are rounded.

vowel, and the [ə] is defined as half-rounded instead of not-rounded. In the IPA vowel quadrilateral we interpreted the [æ] and [ɐ] as not rounded, the [ə] as half rounded and the [ʊ] as rounded.

In the original system, all the possible lengths were not available for all the vowels. In our adapted system for vowels, all lengths are allowed. The correct use of length marks is assumed to be the responsibility of the transcriber. In the original system, for short sounds the feature *long* is set to 3, for half-long sounds to 2 and for long sounds to 1. We have reversed this order: for short sounds the feature *long* gets the value 1, for half-long sounds the value 2 and for long sounds the value 3.

Besides the features given by Vieregge et al., some extra features are added. To be able to process some diacritics, for both the RND data and the NOS data we added the feature *nasal*. For the NOS data we also needed to add the features *breathy*, *creaky*, *toneme 1*, *toneme 2* and *circumflex*.

A feature *vowel* was also added. Usually for vowels this feature is set to 1 and for consonants this feature is set to 0. However, for the [j] and [w] the feature is set to 1 as well. In our system the [i], [j], [u] and [w] are defined as both vowels and consonants. The [j] and the [w] share all the vowel features of respectively the [i] and the [u], and the [i] and the [u] share all the consonant features of the [j] and the [w] (see Section 3.1.3.2). When counting frequencies, both the vowel features and the consonant features are counted for these sounds, however, they are weighted by half. When finding the distance between two segments which are defined as both vowel and consonant, first the distance on the basis of the vowel features is calculated, and next the distance on the basis of the consonant features is found. The final distance between the two segments is equal to the mean of vowel distance and the consonant distance.

The feature *vowel* also plays a role in the definition of silence; a definition of silence in terms of vowel features will be used in the Levenshtein algorithm (see Section 5.1). We defined it to be equal to the schwa, except that the feature *vowel* is set to 0.

	front		central		back	
close	i	y	ɨ	ʉ	ɯ	u
near-close	ɪ	ʏ				ʊ
close-mid	e	ø	ɘ	ɵ	ɤ	o
central			ə			
open-mid	ɛ	œ	ɜ	ɞ	ʌ	ɔ
near-open	æ		ɐ			
open	a	ɶ			ɑ	ɒ

Table 3.8: The IPA vowels defined using the features of Vieregge et al. by analogy with the IPA vowel quadrilateral. Elements left in a cell are spread, and elements right are rounded.

3.1.3.2 Consonants

For getting perceptual distances between 18 Dutch consonants Vieregge et al. performed a perception experiment in which 25 first year speech therapy students were presented with pairs of consonants in medial word position. They were asked to rate each pair on articulatory dissimilarity on a scale from 1 (minimal dissimilarity) to 10 (maximal dissimilarity). The stimulus material consisted of $(18^2 - 18)/2 = 153$ word pairs which differed as little as possible. The stimuli were offered in random order on paper.

Next, a feature system was developed to model perceptual distances as accurately as possible. Features for both place and manner of articulation can be found, comparable to the IPA system, as well as a feature for distinguishing between voiced and voiceless consonants.

The system was originally developed for Dutch, where only for a subset of the Dutch consonants the perceptual distances were measured, viz., the [p], [b], [t], [d], [k], [f], [v], [s], [z], [x], [m], [n], [ŋ], [l], [ʀ], [w], [j] and [h]. Along the lines of Vieregge et al., Cucchiarini (1993) extended the system so that it can be used for Dutch, Limburg and Czech. She replaced the feature *flap* by the feature *trill*. Along the lines of Vieregge et al. she added a number of consonants. For consonants we used the system of Cucchiarini as a basis. Along the lines of Cucchiarini's system we introduced extensions so that it contains all pulmonic consonants of the IPA system. The possible values for the features are listed in Table 3.9. The relation between the manner of articulation (IPA columns) and the relevant features of Cucchiarini is given in Table 3.10. However, although palatal consonants are defined as distributed, the [j] is defined as non-distributed. The relation between the place of articulation (IPA rows) and the relevant Cucchiarini features is given in Table 3.11. The feature *voice* is defined exactly as in the IPA consonant table.

Feature	Value	Meaning
consonant	0	no
	1	yes
place	1.0	bilabial/labiodental
	1.5	dental
	2.0	alveolar/postalveolar
	2.5	retroflex
	3.0	palatal
	4.0	velar/uvular
	4.5	pharyngeal
	5.0	glottal
voice	0	voiceless
	1	voiced
nasal	0.0	not nasal
	0.5	half-nasal
	1.0	nasal
stop	0	no
	1	yes
glide	0	no
	1	yes
lateral	0	no
	1	yes
fricative	0	no
	1	yes
trill	0	no
	1	yes
high	0	no
	1	yes
distributed	0	no
	1	yes
syllabic	0	no
	1	yes
apical	0	no
	1	yes

Table 3.9: The consonant features of Cucchiarini and their possible values. We extended the system with some extra features, in this table given in italics. Only the first twelve features in this table are used for the RND data, the last feature is added for the NOS data.

IPA	place	high	distributed
bilabial	1.0	0	1
labiodental	1.0	0	0
dental	1.5	0	0
alveolar	2.0	0	0
postalveolar	2.0	1	1
retroflex	2.5	0	0
palatal	3.0	1	1
velar	4.0	1	0
uvular	4.0	0	0
pharyngeal	4.5	0	0
glottal	5.0	0	0

Table 3.10: Relation between IPA columns (manner of articulation) and the relevant features of Cucchiarini.

Besides the basic features given by Cucchiarini, some features are added to process some diacritics. For both the RND data and the NOS data we added the feature *syllabic*. For the NOS data only we added the feature *apical*.

A feature *consonant* is added. Usually for the vowels this feature is set to 0, and for the consonants this feature is set to 1. However, for the [i] and the [u] the feature is set to 1 as well. In our system the [i], [j], [u] and [w] are defined as both vowels and consonants. The [i] and the [u] share all the consonant features of respectively the [j] and the [w], and the [j] and the [w] share all the vowel features of [i] and [u] (see Section 3.1.3.1 for more details).

The feature *consonant* also plays a role in the definition of silence; a definition of silence in terms of consonant features will be used in the Levenshtein algorithm (see Section 5.1). We defined it to be equal to the glottal stop, except that the feature *consonant* is set to 0.

3.1.4 Features Almeida and Braun

At the same time as the Vieregge system was developed an alternative system with the same goal was developed which was based on the IPA tables. The system was first developed in the phonetics department of the research institute for German language “Deutscher Sprachatlas” (Marburg, Germany) in 1980 and was further developed and formalized later (Almeida and Braun, 1986). In contrast to the Vieregge system the Almeida & Braun system is articulation-based. The system relies on the assumption that transcription is a process which first consists in an imitation of the relevant utterance, followed by an inference on the part of the transcriber of the articulatory gestures of the speaker, and finally in a phonetic

IPA	nasal	stop	glide	lateral	fricative	trill
plosive	0	1	0	0	0	0
nasal	1	0	0	0	0	0
trill	0	0	0	0	0	1
tap or flap	0	0	0	0	0	0
fricative	0	0	0	0	1	0
lat. fric.	0	0	0	1	1	0
approximant	0	0	1	0	0	0
lat. appr.	0	0	0	1	0	0

Table 3.11: Relation between IPA rows (place of articulation) and the relevant features of Cucchiarini.

description thereof (Almeida, 1984; Almeida and Braun, 1985). The description is carried out in terms of the criteria used by the International Phonetic Alphabet (the version revised to 1993) which essentially consists in an abbreviation for a combination of articulatory features.[4] The Almeida & Braun system is an articulatory system in which sound distances are derived from the IPA vowel quadrilateral and the IPA consonant table. From the beginning the system covers the complete IPA vowel and pulmonic consonant set. Furthermore, in the original system a large number of suprasegmentals and diacritics can be processed.

In our research we introduced adjustments to the Almeida & Braun system. The description given in this section is based on Heeringa and Braun (2003). We describe first the definitions of the vowels and next we explain how the definitions for the consonants were determined.

3.1.4.1 Vowels

The basis for finding vowel distances is the IPA vowel quadrilateral as given in Appendix A Figure A.2. The quadrilateral reflects three features: *advancement*, *height* and *roundedness*. The possible values for the features are listed in Table 3.12. In the vowel quadrilateral we regard the distance between e.g. ɛ vs. ɜ (advancement: front vs. central), ɛ vs. æ (height: open-mid vs. open), and ɛ vs. œ (rounded: no vs. yes) as one step. So when simply subtracting the corresponding feature values from each other and taking the absolute value, we get a distance of one for each of these three pairs.

In the IPA vowel quadrilateral we interpreted the [æ] and [ɐ] as not rounded, the [ə] as half rounded and the [ʊ] as rounded.

[4]The system can be found in the *Handbook of the International Phonetic Association* (IPA, 1999) as well as via: `http://www2.arts.gla.ac.uk/IPA/ipachart.html`.

Feature	Value	Meaning
vowel	0	no
	1	yes
advancement	1	front
	2	central
	3	back
height	1	close
	2	near-close
	3	close-mid
	4	central
	5	open-mid
	6	near-open
	7	open
roundedness	0	no
	1	yes
long	0.0	short
	0.5	half-long
	1.0	long
nasal	0.0	not nasal
	0.5	half-nasal
	1.0	nasal
breathy	0	no
	1	yes
creaky	0	no
	1	yes
toneme 1	0	no
	1	yes
toneme 2	0	no
	1	yes
circumflex	0	no
	1	yes

Table 3.12: The vowel features of Almeida and Braun and their possible values. We extended the system with some extra features, in this table given in italics. Only the first seven features in this table are used for the RND data, the last five features are added for the NOS data.

Besides the basic features derived from the vowel quadrilateral, some extra features are added. To be able to process some suprasegmentals and diacritics, we added the features *long* and *nasal* for both the RND data and the NOS data. Only for the NOS data we added the features *breathy*, *creaky*, *toneme 1*, *toneme 2* and *circumflex*.

A feature *vowel* was also added. Usually for the vowels this feature is set to 1 and for the consonants this feature is set to 0. However, for the [j] and [w] the feature is set to 1 as well. In our system the [i], [j], [u] and [w] are defined as both vowels and consonants. The way in which these sounds are defined and processed is similar as in the V & C system (see Section 3.1.3.1 for more details).

The feature *vowel* also plays a role in the definition of silence. A definition of silence in terms of vowel features will be used in the Levenshtein algorithm (see Section 5.1). We defined it to be equal to the schwa, except that the feature *vowel* is set to 0.

3.1.4.2 Consonants

In our system we only use the pulmonic consonants, the non-pulmonic ones are not included. The basis for finding consonant distances is the IPA table for pulmonic consonants as given in Appendix A Figure A.4. In this table it can be seen that in our system the voiced labial-velar approximant [w] is regarded as, and will be treated as, a bilabial approximant.

The table reflects three features: *place*, *manner* and *voice*. We regard both *place* and *manner* as a scale. The feature *place* gives the *location* of closure and ranges from front to back. The feature *manner* gives the *degree* of closure with roughly the following degrees: complete closure (plosives), oral closure (nasals), intermittant closure (trills, tap and flap), friction (fricatives) and frictionless approximation (approximants). The possible values for the features are listed in Table 3.13. In the consonant table we regard the distance between e.g. [z] vs. [ɾ] (manner: fricative vs. tap or flap), [z] vs. [ʒ] (place: alveolar vs. postalveolar) and [z] vs. [s] (voice: voiced vs. voiceless) as one step. So when simply subtracting the corresponding feature values from each other and taking the absolute value, we get a distance of one for each of these three pairs. We regard the distance between e.g. [ɱ] and [v] (manner: fricative vs. approximant) and [ʙ] and [r] (place: bilabial vs. alveolar) as two steps, although they may be regarded as neighbors.

Besides the basic features derived from the consonant table, some features are added to process some diacritics. For both the RND data and the NOS data we added the feature *syllabic*. For the NOS data only we added the feature *apical*.

A feature *consonant* is added. Usually for the vowels this feature is set to 0, and for the consonants this feature is set to 1. However, for the [i] and the [u] the feature is set to 1 as well. In our system the [i], [j], [u] and [w] are defined

Feature	Value	Meaning
consonant	0	no
	1	yes
place	1	bilabial
	2	labiodental
	3	dental
	4	alveolar
	5	postalveolar
	6	retroflex
	7	palatal
	8	velar
	9	uvular
	10	pharyngeal
	11	glottal
manner	1	plosive
	2	nasal
	3	trill
	4	tap or flap
	5	fricative
	6	lateral fricative
	7	approximant
	8	lateral approximant
voice	0	no
	1	yes
syllabic	0	no
	1	yes
apical	0	no
	1	yes

Table 3.13: The consonant features of Almeida and Braun and their possible values. We extended the system with some extra features, in this table given in italics. Only the first five features in this table are used for the RND data, the last feature is added for the NOS data.

as both vowels and consonants. The way in which these sounds are defined and processed is similar as in the V & C system (see Section 3.1.3.2).

A definition of silence in terms of consonant features will be used in the Levenshtein algorithm (see Section 5.1). We defined it to be equal to the glottal stop, except that the feature *consonant* is set to 0.

3.2 Diphthongs

A diphthong is a vowel with a changing color. In the feature table of H & H diphthongs are combinations of two vowels, where the first segment is short or half-long, and the second short. When the first element is long, two succeeding vowels are not regarded as a diphthong.

When processing diphthongs, we need an accurate representation of them. Vieregge et al. write that they follow Moulton (1962) and consider a diphthong as vowel+vowel sequence, the second vowel being non-syllabic allophonically. On the other hand, H & H regard regular diphthongs as segmental units (Hoppenbrouwers and Hoppenbrouwers, 2001, p. 35). We make no a priori decision here but experimented with both two-phone and one-phone representations. Validation work will make clear whether the different representations result in different results, and which representation gives the better results (see Chapter 7).

In this section we explain how diphthongs can be defined as segmental units using the different segment representations. An accurate representation should reflect the color at the onset, the transition, and the color of the offset. The second element as noted in a transcription may not always be the real offset. It may also be a target position which is not really pronounced (Rietveld and Van Heuven, 1997, p. 74). However, with the discrete representations we used it is not possible to represent diphthongs in such a refined way. E.g. the way in which the transition takes place cannot be represented. In our research we used a simplied approach in which the definition of a diphthong is based only on the color of the onset and the color of the target position.

When using the phone representation, all sounds are equally different. Therefore, when a diphthong is regarded as one segmental unit, e.g. the [a] and [au] will be regarded as equally different as the [a] and the [i], except that the [au] (like all diphthongs) is treated as a long sound. So no special specifications for diphthongs need to be made. This is a very rough approach where the color of the onset and the target position plays no role. However, when using feature representations, the definition of diphthongs is based on the definitions of the onset color and the target color. Below we explain how diphthongs are specified in the feature system of H & H and how we defined them for the systems of V & C and A & B.

H & H make a distinction between *closing diphthongs* and *centering diphthongs*. A closing diphthong is a long vowel with a movement toward a non-central position in the vowel space. On the contrary a centering diphthong is a vowel with a movement toward a central position in the vowel space, the schwa. In Section 3.2.1 we describe the specification of closing diphthongs, and in Section 3.2.2 we explain the specification of centering diphthongs.

3.2.1 Closing diphthongs

As mentioned above a closing diphthong is a long vowel with a movement toward a non-central position in the vowel space. The term *closing* indicates that there is a movement in the direction of a closer vowel. This movement can be vertical (e.g. [ɔu]) or diagonal (e.g. [ɔi]). In a diphthong, the color of a sound changes from a start position to an end position. Therefore, when specifying a diphthong the feature values are derived from the feature values of the vowel corresponding with the start position and the vowel corresponding with the end position. If the second element of a diphthong is noted as a [j] or a [w], we used the feature values of the [i] and [u] respectively. In the system of H & H the feature values of closing diphthongs are defined as follows:

front	:	mean of both segments
back	:	mean of both segments
round	:	mean of both segments
low	:	value of first element
polar	:	value of first element
long	:	always [+long]
peripheral	:	always [+peripheral]
diphthong	:	always [+diphthong]

A movement from front to back (or vice versa) is specified by using the mean of the start position and the end position. H & H symbolize the mean value using '$*$'. It is striking that a similar procedure is not followed with respect to the height. The features *low* and *polar* simply get the value of the first segment. Does this mean that the first segment is most dominant? H & H write that this vertical closing movement is specified by specifying the feature *diphthong* as positive (Hoppenbrouwers and Hoppenbrouwers, 2001, p. 46). Because there is always a movement to either the [i] or [u] in closing diphthongs, they are specified as [+peripheral] which makes closing diphthongs a bit more related to the (half-)long [i] and [u], which are also specified as [+peripheral] (see Table 3.2). Closing diphthongs are specified as peripheral sounds regardless whether the (half-)long version of the first element is specified as peripheral or not.

For the system of V & C we define the feature values of closing diphthongs as follows:

advancement	:	mean of both segments
high	:	mean of both segments
long	:	always *long*=3
rounded	:	mean of both segments

In contrast to H & H the value of *high* is determined in the same way as the value for *advancement*. In our opinion the value for *high* should also be based on both segments. A disadvantage of this approach compared to that of H & H is that the order of segments is not represented. For example the [ɑu] and the [uɑ] are defined in exactly the same way. For both the RND and the NOS data this is no problem since in the selection of diphthongs in both data sources none of the diphthongs has a symmetric counterpart. However, when a selection of diphthongs contains symmetric cases, a solution may be to weight the advancement, height and rounding of the first element 75%, and the same features of the second element 25%.

For the system of A & B we define the feature values of closing diphthongs as follows:

advancement	:	mean of both segments
height	:	mean of both segments
roundedness	:	mean of both segments
long	:	always *long*=1

Just as in the V & C system for the feature *high* in the A & B system the value for the feature *height* is based on both segments, which is different from the approach of H & H. Again symmetric diphthongs cannot be distinguished from each other when using our definition. As noted above, however, for both the RND and the NOS data no symmetric cases were selected.

Using the RND data we adopted the selection of closing diphthongs as made by H & H. In Table 3.14 the closing diphthongs which were included in the feature table of H & H are listed, extended with six diphthongs which are lacking in the H & H table. The fact that these six diphthongs were missing has to do with the fact that the monophthong [ɒ] is also missing from the table.

When a closing diphthong is noted as a combination of vowel+vowel, for some diphthongs the first element is short, for others the first element is half long. The first element is never long. If a closing diphthong is noted as a combination of vowel+consonant, then the first element is always short. The first element is never half-long or long. According to this H & H write that only sequential diphthongs of the type [aːj], consisting of a long vowel followed by a [i] or [j] as in Dutch *fraai* and *mooi* are biphonemically interpreted (Hoppenbrouwers and Hoppenbrouwers, 2001, p. 35). For all closing diphthongs the second element should be short.

Example		Diphthong	Defined as
		[ʏy]	[œy]
		[ʊi]	
Dutch:	hooi	[ʊˑi]	
English:	bay	[ei]	
		[eˑi]	
		[øi]	[øy]
		[øy]	
		[oj]	[ʊj]
		[ou]	
		[oˑu]	
		[ɛi]	
Dutch:	wijn	[ɛˑi]	
		[ɛj]	
		[œi]	[œy]
		[œy]	
Dutch:	huis	[œˑy]	
English:	boy	[ɔi]	
		[ɔˑi]	
		[ɔj]	
		[ɔu]	
Dutch:	koud	[ɔˑu]	
		[ɔw]	
		[æi]	
		[æˑi]	
		[æj]	
English:	line	[ai]	[æi]
		[ay]	[ɶy]
		[aˑy]	[ɶy]
		[ɑi]	
German:	drei	[ɑˑi]	
		[ɑj]	
		[ɑu]	
Dutch:	saus	[ɑˑu]	
		[ɑw]	
		[ɒi]	
		[ɒˑi]	
		[ɒj]	
		[ɒu]	
		[ɒˑu]	
		[ɒw]	

Table 3.14: Dutch closing diphthongs as found in the feature table of H & H. The last six diphthongs were not originally included in the table.

The selection of closing diphthongs for the NOS data is much smaller. In the Norwegian text ‘Nordavinden og sola’ the only probable diphthong is the *ei* as in *dei* ‘them’, *ein* ‘a’, *seg* ‘him, her’, *dei* ‘they’, *einige* ‘agreed’, *skein* ‘shone’ and *meir* ’more’. In the cases where the ‘potential diphthongs’ in these words were perceived, one of the following six transcriptions could be used: [ei], [eɪ], [ɛi], [ɛɪ], [æi] and [æɪ]. Both the first and the second element are short. All suprasegmentals and diacritics which may be applied to monophthongs may also be applied to these diphthongs.

In the H & H table no closing diphthongs are defined with suprasegmentals and diacritics (except for length as just explained). In our research all suprasegmentals and diacritics which we allow to be applied to monophthongs (see Section 3.4) may also be applied to closing diphthongs. When a suprasegmental or diacritic is noted after the first or second segment of a diphthong, it is applied to the diphthong as a whole.

In the RND the second element of a closing diphthong is often noted as extra-short (i.e. in superscript or with a smaller character), probably with the goal to express that the first element is more significant than the second element. In the table of H & H the second element of the diphthong is always noted in normal script. However, when it was noted as extra-short, we treated it as a short sound. In the NOS data a second element of a closing diphthong noted as extra-short was never observed. However, if this should occur, it will be treated as short as well, just as for the RND data.

3.2.2 Centering diphthongs

A centering diphthong is a vowel with a movement toward a central position in the vowel space, the schwa. So when specifying a centering diphthong, the feature values should be derived from the the feature values of the first vowel and the ending schwa. In the feature system of H & H the feature values of centering diphthongs are defined as follows:

front	:	mean of both segments
back	:	mean of both segments
round	:	mean of both segments
low	:	mean of both segments
polar	:	mean of both segments
long	:	always [+long]
peripheral	:	mean of both segments
diphthong	:	always [−diphthong]

A movement from front to back (or vice versa) is defined by taking the mean of the start position and the end position. Different from the definition for closing

Example		Diphthong
Afrikaans:	tee	[iɘ]
		[yɘ]
Afrikaans:	voet	[uɘ]
Frisian:	each	[ɪɘ]
		[ʏɘ]
		[ʊɘ]
		[eɘ]
Afrikaans:	deur	[øɘ]
Frisian:	roas	[oɘ]
		[ɛɘ]
		[œɘ]
		[ɔɘ]
		[æɘ]
		[aɘ]
		[ɑɘ]
		[ɒɘ]

Table 3.15: Dutch centering diphthongs as found in the feature table of H & H. The last diphthong was not originally included in the table.

diphthongs, a movement from low to high (or vice versa) is also defined by taking the mean of the start position and the end position. Centering diphthongs also differ from closing diphthongs in the way the feature *peripheral* is defined. For centering diphthongs the mean of both elements is used. For finding this value as first element the (half-)long sound is used, which may be specified as either [−peripheral] or [+peripheral] (see Section 3.1.2.1).

The feature values for the centering diphthongs in the systems of V & C and A & B are found in the same way as for the closing diphthongs (see Section 3.2.1) which is analogous to the way in which H & H find the feature values for centering diphthongs.

Using the RND data we adopted the selection of centering diphthongs as made by H & H, just as we did for the closing diphthongs. In Table 3.15 the centering diphthongs which were included in the feature table of H & H are listed, extended with one diphthong which is missing in the H & H table, due to the monophthong [ɒ]'s being missing.

In the combination vowel+vowel the first element is always short. The first element is never half-long or long. The second element is always short as well.

For the NOS data no centering diphthongs are selected. In the Norwegian text 'Nordavinden og sola' the only probable diphthong is the [iə] as in [^{2}eːniə] *enige* 'agreed'. This pronunciation (or something similar) was found in the dialects of Bø, Borre, Larvik, Stavanger and Trysil. Here the [iə] may be an shortened form of [igə] since we found in the dialect of Fyresdal the pronunciation [²eːnigə]. Therefore, we decided not to include the sequence [iə] as centering diphthong.

In the H & H table for all centering diphthongs a nasalized version is also defined. In our research all suprasegmentals and diacritics which we allow to be applied to monophthongs (see Section 3.4) may also be applied to centering diphthongs, just as for closing diphthongs. When a suprasegmental or diacritic is noted after the first or second segment of a diphthong, it is applied to the diphthong as a whole.

In the RND the second element of a centering diphthong is often noted as extra-short, just as for closing diphthongs (see Section 3.2.1). However, when it was noted as extra-short, we treated it as a short sound. For the NOS no centering diphthongs were selected. However, if centering diphthongs would be processed and the second element is noted as extra-short, the second element will be treated as short as well, just as for the RND data.

3.3 Affricates

In the RND data no affricates are used. However, in the NOS data they do appear. When processing them, both elements are processed as extra-short independent elements in sequence. E.g. [t͡s] is treated as [t̆s̆]. The idea behind this is that two extra-short elements are one segment of normal length together. This can be illustrated with an example. The Standard German word for 'plough' is *pflügen.* In Low-German dialects the word *plögen* (or something similar) is often used. However, in a small southerly part of this northern area the words *flichen* and *flien* are used, while in the Prussian area in the northwest the word *flieje* (besides *pleje*) is used (König and Paul, 1991, p. 198). Here we see that the [p], [pf] and [f] correspond to each other. Regarding a [pf] as a sequence of an extra-short [p] and an extra-short [f] allows us to match both elements with one segment of normal length, either a [p] or [f].

In some cases the first element of an affricate is stressed more than the second, or the second element is stressed more than the first. E.g. in the Sardinian dialect of Atzara the equivalent for 'daughters' is pronounced as [ˈfid͡ʒaza]. In the affricate [d͡ʒ] the first element should be processed as a short sound, and the second as an extra-short sound. In the Sardinian dialect of Abbasanta the equivalent for 'policeman' is pronounced as [polit͡sɔtˑɔ]. In the affricate [t͡s] the first element should be processed as an extra-short sound, and the second as a short sound. In our research these cases should not be noted as affricates, but as sequences of a short sound followed by an extra-short sound and an extra-short sound followed by a short sound respectively.

3.4 Suprasegmentals and diacritics

Using different sound representations it is possible to process suprasegmentals and diacritics. We did not implement the processing of all suprasegmentals and diacritics, so in this section only a subset is examined. The selection of suprasegmentals and diacritics was determined by the fact that they appear in the transcriptions we used on the one hand, and by the possibilities of the sound representations on the other hand.

When processing suprasegmentals and diacritics it is important to find and use the right weights which represent as precisely as possible the effect as perceived by listeners. In the next sections, all weights proposed are not based on real measurements. They are intuitively assigned. Besides, for different language groups different weights should be used. Our starting point is mainly the Dutch language area.

3.4.1 Stress and tonemes

In the RND, one symbol is available for marking stress [ˈ] which we interpret as corresponding with primary stress. The RND does not consistently mark which syllable is stressed for every word. It may be that stress is only noted when a syllable is stressed that differs from the one which the transcriber expected to be stressed.

In most Dutch dialects tonemes play no role. Only in the Limburg dialects can tonemes be found. In the RND, the dialects of the Limburg area were recorded by four transcribers. The transcribers did not note tonemes in equal detail (see part 8 of the RND). The fact that stress and tonemes are not consequently noted in the RND material may be the reason why these are not processed in the H & H system. For the RND, we also did not process these suprasegmentals.

In the IPA system we find symbols for primary stress [ˈ] and secondary stress [ˌ]. However, most Norwegian dialects are pitch accent varieties. All syllables with primary stress generate tone, or 'accent' (or 'tonal accent'). Also: tonal accent can only be generated from primary stressed syllables. Only a few Norwegian dialects lack the tonal/accentual opposition. They generate the same tone, or accent, for all words, which is the same as primary stress in the IPA system. These dialects are found in an area around Bergen, in the Brønnøy area north of Trondheim and in many dialects of the two northernmost counties, Troms and Finnmark. In the varieties with tonal/accentual opposition three tonemes may occur: toneme 1 and toneme 2 (Kristoffersen, 2000) and circumflex (Almberg, 2001). Since no symbols are available in the IPA system for these tonemes, extra symbols are introduced and used in the NOS transcriptions. Syllables with toneme 1 are preceeded by a [ˈ], syllables with toneme 2 by a [^{2}], and syllables with circumflex by a [˜].[5] All transcriptions of the NOS data were made by one transcriber, who noted stress and tonemes consistently. Therefore, we process stress and tonemes for the NOS data when using feature representations. For the phone representation we found no way to deal with them.

In the transcriptions stress and toneme marks are noted before a syllable. To be able to process these marks, we shift them to the first vowel in the syllable.[6] So stress and tonemes are processed like properties (features) of a vowel. We suppose that stress and tonemes are mainly realized by the way in which vowels in syllables are pronounced. When diphthongs are processed as one sound and a stress or toneme mark is noted before the second element of a diphthong, it is shifted before and applied to the first vowel to the right. This may happen when the first element is the last segment of the one syllable, and the second element is the first segment of the next syllable.

To be able to process stress and tonemes as properties of vowels, we extended the feature systems of H & H, V & C and A & B with three features: *toneme 1*, *toneme 2* and *circumflex*. With these features the different stresses and tonemes are represented as follows:

	toneme 1	toneme 2	circumflex
primary stress	0.250	0.250	0.250
secondary stress	0.125	0.125	0.125
stress and toneme 1	0.500	0.250	0.250
stress and toneme 2	0.250	0.500	0.250
circumflex	0.250	0.250	0.500

On the basis of these representations, the distances between the stresses and tonemes can be calculated as the sum of the differences per feature (see Section 3.6.2 for more explanation and other alternatives). This results in the following distances:

	nothing	secondary	primary	toneme 1	toneme 2	circumflex
nothing		0.375	0.750	1.000	1.000	1.000
secondary			0.375	0.625	0.625	0.625
primary				0.250	0.250	0.250
toneme 1					0.500	0.500
toneme 2						0.500
circumflex						

The scheme reflects the view that primary stress weighs more heavily than secondary stress. The three tonemes are regarded as equally different from one

[5]The reader might expect that the circumflex should be noted by a [ˆ], but this symbol is reserved for a tone with a falling contour in the IPA system.

[6]Here the [w] and the [j] are considered as consonants.

another. In our scheme this distance is equal to 0.500. Because all tonemes imply primary stress, the distance between the toneme of an accent variety and the primary stress of a non-accent variety should not be too large. Therefore, in the scheme the relative small distance of 0.250 is found. With respect to 'nothing' the tonemes weigh a little bit more than primary stress only: 1.000 versus 0.750.

3.4.2 Quantity

In this section we discuss the processing of quantity marks: extra-short, half-long, long and syllabic. In the RND extra-short sounds are noted in superscript or with a smaller character. In the IPA system extra-short sounds are noted with a ˘ on top of the sound symbol. In both the RND and the IPA system half-long sounds are followed by a ˑ, long sounds are followed by a ː.

3.4.2.1 Extra-short

H & H process extra-short sounds in two ways. First, for some few scattered cases in Groningen dialects and eastern and western Flemish dialects the odd cases in the transcriptions were taken into account while the even cases were ignored. Second the feature table was extended with specifications for the extra-short versions of the [r], [ɣ], [m], [n], [ŋ] and [h]. For these sounds both the odd and even cases are processed, using the specifications of the extra-short versions as given in the feature table. In the specifications the values of non-redundant positively marked features are halved.

It was not clear to us why only a restricted set of sounds may be processed as extra-short. In our research, all sounds may be processed as extra-short. Furthermore, the way in which H & H process extra-short sounds works for feature-based comparison methods, but not for phone-based methods. Therefore, we used another way of processing them which works for phone-based representations as well, and gives the same effect as the approach of H & H when using the feature frequency method. In our approach the half weighting of an extra-short sound with respect to other sounds is realized by changing the transcription beforehand. We retain the extra-short sounds as they are and double all other sounds. E.g. the Dutch word *arm* 'arm' is sometimes pronounced as [ɑrə̆m]. This word is processed as [ɑɑrrəmm]. The Dutch word *timmerman* 'carpenter' is sometimes pronounced as [tɪmər̆mɑn]. This word is changed in [ttɪɪmməərmmɑɑnn].

3.4.2.2 Half-long, long (1)

In the feature table of H & H all *monophthongs* are defined for three lengths: short, half-long and long. In the table half-long and long vowels are not really distinguished. In fact half-long vowels are processed as long vowels. Both half-long and long vowels are defined as [+long] (and [+peripheral] for peripheral vowels). For the RND data this is a sound approach to eliminate the influence of the different uses of length marks per transcriber or per atlas part. So for the RND data we follow H & H when deriving feature representations for these sounds.

However, for the NOS data we do distinguish half-long sounds from long sounds. Using the system of H & H half-long NOS vowels are symbolized as [∗long] (and [∗peripheral] for peripheral vowels). We recall that the '∗' is used to signify an intermediate value. In the system of V & C the feature *long* is set to 2, in the system of A & B to 0.5. Half-long RND vowels and long RND and NOS vowels are specified as [+long] (and [+peripheral] for peripheral vowels) in the system of H & H. In the system of V & C the feature *long* is set to 3, in the system of A & B to 1. Note that in the system of V & C length is weighted more heavily than in the two other feature systems.

For the [ɘ], [ʉ], [ɵ] and [ɜ] only versions without length marks are defined in the feature table of H & H. Furthermore, in Section 3.1.3 we saw that in the original V & C system not all vowels can have all lengths. In our adapted systems, for both the RND and the NOS for all monophthongs all lengths are allowed and processed in the way we explained above.

In the feature system of H & H "length" is also processed for a restricted number of *consonants*. While in the RND the nasals ([n], [m], [ŋ]) are noted as long, H & H define them as syllabic sounds by specifying them as [+vowel] and [+syllabic].[7] Long nasals are always treated as syllabic, regardless of the context in which they appear, following H & H. When using the systems of V & C and A & B, we treat such nasals as syllabic sounds as well. In the two systems the feature *syllabic* is set to 1. For half-long nasals H & H ignore the length mark, possible because they judge a half-long nasal as too similar to a short nasal. Here again we follow H & H. For other consonants, H & H did not process length marks. For Dutch varieties it is not common to lengthen other consonants. In the RND some half-long non-nasal consonants can be found when they simultaneously form the last segment of the one word and the first segment of the next word. Since these are rare cases on the one hand, and as we limit our study to single words on the other hand, we also ignore length symbols of non-nasal consonants.

Although it is justified for the RND to interpret a long nasal as a syllabic sound, in a more general approach long nasals and syllabic nasals should be distinguished. In the NOS data we found that *the northwind* (nordavinden) was pronounced as [²nuːɾɑˌʋinːn̩] in the dialect of Oslo. This makes clear that long nasals and syllabic nasal are not the same. Therefore, for the NOS data we did not process long nasals as syllabic, but expect that the transcriber has put a syllabic mark [̩] under a sound if it should be interpreted as syllabic. The problem how to process (half-)long nasal and non-nasal consonants remains, for

[7]In the feature table of H & H syllabic consonants are specified as [+vowel] *and* [+consonant]. Syllabicity makes consonants more vowel like.

example geminates. For both vowels and consonants length refers to duration. Even so, we have the feeling that the feature which represents vowel length should not be used for consonant length. Therefore, for the NOS data consonant length is not processed, unless an author transcribes for example a [tː] as [tt]. We did not find this type of notations in the NOS data.

As we saw, when using a feature-based representation, length can easily be processed by changing one or more feature values. However, this does not work when using the phone-based representation. In some languages length is redundant to some extent. E.g., in Standard Dutch the [e], [a] and [o] (written as <ee>, <aa> and <oo> respectively in closed syllables) are usually long, while the [ɛ], [ɑ] and [ɔ] (written as <e>, <a> and <o> respectively in closed syllables) are short. Therefore, for the phone-based representation we experimented with an approach in which half-long and long are not processed.

3.4.2.3 Half-long, long (2)

Although half-long and long may sometimes be redundant to some extent, this will never consistently be the case. When ignoring both length marks, there is no difference between e.g., the Dutch word *ver* [fɛr] 'far' and *fair* [fɛːr] 'fair'. In this section we present a second approach for processing *half-long* and *long* which we examined in addition to the approach which we described in the previous section. The benefit of this approach is that *half-long* and *long* can be processed not only when using a feature representation, but also when using the phone representation as well.

In Zwaardemaker and Eykman (1928) it was found that the duration of short vowels in Dutch is 40-50% of that of long vowels (p. 298). In a study of durational properties of vowels in Dutch, Nooteboom (1972) found that the duration of long vowels is about two times the duration of short vowels (p. 115). However, this ratio may be affected by stress, the position within the word, position within the sentence, speech rate, etc. For Norwegian Fintoft (1961) found that the duration of short vowels is 53% of that of long vowels (p. 24). This ratio is based on a number of nonsense words built up on the structural principles of real Norwegian words. The words were read by Norwegian speakers. More ratios of short to long vowels for different languages can be found in Elert (1964).

Although the duration ratios may vary per language and under different conditions, we take as a starting hypothesis that the duration of long vowels is twice the duration of short vowels. The duration of half-long vowels is intermediate between the duration of short and long vowels. Analogous to vowels, for short, half-long and long consonants the same ratios are taken as starting point. This implies that a long consonant is processed as if it were just as long as a long vowel. In a more refined system vowel-consonant ratios should also be taken into account. In the present system the different durations are implemented by changing the transcription. Above we explained that extra-short sounds are retained

```
if extra-short
  then retain sound
  else if normal
         then double sound
         else if half-long
                then treble sound
                else if long
                       then quadruple sound
                       else{nothing}
```

Figure 3.1: Procedure followed when more than one length mark is noted for the same phone.

as they are, and short sounds are doubled. In this outline we treble half-long sounds, and quadruple long sounds. In this approach, the suprasegmentals *half-long* and *long* are processed for all vowels and all consonants. E.g. the Dutch word *ook* ‘ook’ is pronounced as [oːk]. This is changed into [ooookk]. In the Sardinian dialect of Abbasanta the word for ‘water’ is pronounced as [abˑa]. This becomes [aabbbaa]. In the same dialect the equivalent for ‘then’ is pronounced as [asːɔɾa]. This is changed in [aassssɔɔɾɾaa].

We applied this approach to both the phone-based and feature-based representations. We are aware of the fact that length is heavily weighted in this procedure, but judge that length plays a rather strong role in perception. E.g. it is for a listener striking when a speaker lengthens vowels at positions where the listener himself would not.

For those cases where a transcriber unfortunately noted more than one length mark for one phone, we follow the procedure as given in Figure 3.1.

3.4.2.4 Syllabic

In the RND consonants may also be vocalized. We process vocalized sounds as syllabic sounds. Vocalized (RND) or syllabic sounds (NOS) are marked with the diacritic *syllabic*. We are aware of the fact that there is no agreed phonetic definition for syllabicity. However, syllabicity forms part of the descriptive framework of the IPA and thus we have to decide how to deal with it since it occurs in both the RND and the NOS transcriptions. We consider two approaches for processing syllabic sounds which corresponds with the two approaches which are regarded when processing *half-long* and *long*.

In the first approach *syllabic* is not processed at all when using the phone representation. This is potentially interesting since *syllabic* may be redundant. E.g. a nasal after a stop at the end of a syllable can hardly be pronounced as a non-syllabic consonant. However, when using feature systems, this diacritic can easily be processed by changing feature values. In the system of H & H a syllabic sound is specified as [+vowel] and [+syllabic]. In the system of both V & C and A & B the feature *syllabic* is set to 1.

In the H & H feature table only for the [m], [n], [ŋ], [r] and [l] syllabic versions are specified. For the RND data we retained the same restriction in the use of the diacritic *syllabic*. For the NOS data we do not check whether the transcriber noted the syllabic diacritic [̩] under a nasal, trill or lateral approximant but assume a correct use of this diacritic by the transcriber. For the NOS data the same way of processing is followed as for the RND.

In the second approach *syllabic* is processed by changing the transcription. In this way *syllabic* is processed for both the phone and the feature representation. Syllabic sounds are processed as long sounds, i.e. they are quadrupled. Here for both the RND and the NOS data the correct use of this diacritic is not checked, but is assumed to be the responsibility of the transcriber.

3.4.3 Place of articulation

3.4.3.1 Advanced, retracted

In the RND vowels and consonants can be followed by a [̙] or a [̘], which means respectively ‘more to the back’ and ‘more to the front’. The same diacritical marks are found in the IPA system, but there they represent respectively the diacritics *advanced tongue root* and *retracted tongue root*. Following H & H, we did not process these diacritics for the RND data, since their use is probably too transcriber-dependent. For the NOS data these diacritics were ignored as well. It was not clear how these diacritics should be processed. However, in the NOS data the diacritical marks [̟] and [̠] also appear, representing respectively the diacritics *advanced* and *retracted*. We processed them for vowels only. We only found a satisfying way to process them in the systems of V & C and A & B.

Using V & C the feature *advancement* is decreased by 1 for an advanced vowel and increased by 1 for a retracted vowel. For the A & B system the feature *advancement* is decreased by 0.5 for an advanced vowel and increased by 0.5 for a retracted vowel. The different weighting of V & C and A & B are due to the different weighting of the feature *advancement*. For both systems the result is that e.g. the [i̠] and the [ɨ̟] are equal, both located exactly in the middle between the [i] and the [ɨ]. Using V & C *advanced* is not processed for vowels with *advancement*=2 (front) while *retracted* is not processed for vowels with *advancement*=6 (back). In the A & B system *advanced* is not processed for vowels with *advancement*=1 (front) while *retracted* is not processed for vowels with *advancement*=3 (back).

3.4.3.2 Raised, lowered

In the RND vowels and consonants can be followed by a [˔] or a [˕], which means respectively 'more closed' and 'more open'. In the IPA system the same diacritical marks are found, representing the diacritics *raised* and *lowered*. Following H & H, we did not process these diacritics for the RND data, since their use is probably too transcriber-dependent. However, for the NOS data they are processed. We processed them for vowels only. We only found a satisfying way to process them in the systems of V & C and A & B.

Using the feature system of V & C the feature *high* is increased by 0.25 for a raised vowel and decreased by 0.25 for a lowered vowel. For the A & B system the feature *height* is decreased by 0.5 for a raised vowel and increased by 0.5 for a lowered vowel. The different weighting for both systems is due to the different weighting of the features *high* and *height*. For both systems the result is that e.g. the [ɛ̞] and the [æ̝] are equal, both located exactly in the middle between the [ɛ] and the [æ]. Using the V & C system *raised* is not processed for vowels with *high*=4 (close) while *lowered* is not processed for vowels with *high*=1 (open). For the system of A & B *raised* is not processed for vowels with *high*=4 (close) while *lowered* is not processed for vowels with *high*=1 (open).

3.4.3.3 Labialized, palatalized, velarized, pharyngealized

In the feature table of H & H the sequence [tj] and the [tʲ] are specified as a [c]. The [c] is only used in part 16 of the RND, therefore, it is obvious to interpret and to process the [tj] and the [tʲ] as substitutes for the [c]. In our research in the RND transcriptions we replaced all [tj]'s and [tʲ]'s by the [c]. In the RND data the diacritic *palatalized* is noted by putting a dot on top of or below the consonant, or by putting a [̥] below the consonant. Except for the [t], as just explained, we did not process this diacritic, following H & H. The use of this diacritic may be too transcriber dependent. In the NOS data the diacritic *palatalized* was not found.

However, for the NOS data the diacritics *labialized* (ʷ), *palatalized* (ʲ), *velarized* (ˠ) and *pharyngealized* (ˤ) are taken into account when using the feature systems of V & C and A & B. They are processed by changing the feature *place*. The new place is based on the bit representation of the original place of articulation and the bit representation of respectively bilabial, palatal, velar and pharyngeal. The original place of articulation is weighted for 75% and the secondary place of articulation for 25%. Since e.g. a velarized [t] is still recognized as a [t] and not as a [k], the original place of articulation should be weighted more heavily then the secondary place of articulation. The weightings are applied to each bit of each pair of bits separately. This assures that the bit representation of the place of a velarized [t] is distinguished from one of the existing places between alveolar and velar (see Section 3.6.2 for a more extended explanation).

3.4.4 Manner of articulation

3.4.4.1 Apical

We found no way to process the diacritic *apical* using phones (e.g. s̺). However, when using a feature system it is possible to process this diacritic. In the RND data this diacritic is not used. However, when processing data based on the modern IPA system such as the NOS data, the occurrence of this diacritic will be taken into account. In the NOS data this diacritic is not found, but in data of e.g. Roman languages this mark may occur. In all three feature systems we added an extra feature *apical*. In the H & H system apical sounds are specified as [∗apical], in the V & C and A & B systems the feature *apical* is set to 0.5.

3.4.4.2 Nasalized

In the RND data sounds can be nasalized (e.g. [ã]), but also half-nasalized (e.g.[ã́]). Half-nasalized sounds may be conceived of being produced with the velum in an intermediate position between fully raised and fully lowered. In the feature table of H & H a nasalized and a half-nasalized version is defined for all monophthongs (possibly for different lengths) and centering diphthongs. In our research both diacritics may also be applied on closing diphthongs and consonants. The use of the diacritic *half-nasalized* is specific for the RND data, in the NOS data only the diacritic *nasalized* is used.

H & H pay special attention to instances of the combination (half-)nasalized vowel + (extra-short) nasal consonant (Hoppenbrouwers and Hoppenbrouwers, 2001, p. 46). All possible combinations are listed in Table 3.16. If we understand the explanation of H & H correctly, then the combinations 1, 3 and 4 are impossible because of undervaluing nasality, and the combinations 6, 9 and 10 are not possible due to overrating nasality. Only the combinations 2, 5, 7 and 8 are possible. H & H corrected the combinations which they judged to be impossible. We are not convinced that all of the combinations are impossible. Therefore, we made no changes in the transcription and process nasality as given by the transcriber.

In the feature table of H & H half-nasalized sounds are specified as [∗nasal], and nasalized sounds as [+nasal]. For both V & C and A & B the feature *nasal* is set to 0.5 for half-nasalized vowels and to 1 for nasalized vowels. Note that for both systems the feature *nasal* is a vowel feature. So with the use of this feature it is not expressed that a nasal consonant is more related to a nasalized vowel than to a non-nasalized vowel, which is a disadvantage of the fact that vowel features and consonant features are separated.

1.	Not nasalized vowel	+	extra short nasal
2.	Not nasalized vowel	+	nasal
3.	Half nasalized vowel	+	nothing
4.	Half nasalized vowel	+	non-nasal
5.	Half nasalized vowel	+	extra-short nasal
6.	Half nasalized vowel	+	nasal
7.	Nasalized vowel	+	nothing
8.	Nasalized vowel	+	non-nasal
9.	Nasalized vowel	+	extra-short nasal
10.	Nasalized vowel	+	nasal

Table 3.16: All possible combinations of a (half-)nasalized vowel and an (extra-short) nasal consonant.

3.4.4.3 No audible release

The diacritic *no audible release* (e.g. [d̚]) appears in the NOS data. However, this diacritic was not processed since it was not clear how it could be processed in the feature system.

3.4.5 Voice

3.4.5.1 Aspiration

H & H did not process the diacritic *aspirated* in their feature system. Possibly the use of this diacritic in the RND is too transcriber-dependent. We follow H & H and ignore this diacritic when processing the RND data. However, for the NOS data source (where all data is transcribed by one transcriber) it is processed. An [h] is inserted after the phone which was noted to be aspirated. This [h] is noted as extra-short, so the weighting is halved. When using feature systems, another way to process *aspirated* would have been to use an extra feature. In that case the weight of aspiration would be much lower. However, our approach reflects the fact that an aspirated sound is perceived as a sound followed by a small [h]. The rather strong weighting accords with the fact that people who aspirate sounds are quickly associated with certain regions.

3.4.5.2 Voiceless, voiced

H & H mention that phonological interpretation systematically seems to play a role in the RND (Hoppenbrouwers and Hoppenbrouwers, 2001, p. 31). In case of assimilation of voice this results in notations like [nit v̥eːl] *niet veel* 'not much'.

In that case H & H replace the voiceless [v] by a [f]. We chose a more cautious approach. When using phones, the diacritics are ignored. However, when using the feature system of H & H, both a voiceless version of a normally voiced consonant (e.g. [v̥]) and a voiced version of a normally voiceless consonant (e.g. [f̬]) are specified as [∗voiced], which means that the consonant is half-voiced and half-voiceless. When using the feature systems of V & C or A & B the feature *voice* is set to 0.5. The procedure for phones and features as described here applies for both, the RND and the NOS data.

3.4.5.3 Breathy, creaky

For the NOS data only the diacritics *breathy voiced* (e.g. [ə̤]) and *creaky voiced* (e.g. [æ̰]) are processed. They are only processed when using a feature representation, since we found no possibility of processing them when using the phone representation. Since these diacritics should not weigh too heavily, we assign only a value of 0.25 to the features *breathy* and *creaky* respectively. These weightings are chosen intuitively. Both diacritics are only processed when noted below a vowel.

3.5 Redundancy

We only find rules that are applied to remove redundancy from feature bundles in the feature system of H & H. This can be explained from the fact that the system was originally developed to be used for the feature frequency method in which features are counted. H & H write that a text which contains many coronals and anteriors, will also contain many consonants, and that in a language with a relatively great number of vowels there will be relatively less space for consonants (Hoppenbrouwers and Hoppenbrouwers, 2001, p. 43). Therefore, a feature specification may contain redundant information. H & H decided to ignore feature values which may be predicted on the basis of other feature values using rules.

In the feature table the presence of a feature is indicated by a '+' and the absence by a '−'. Also a '∗' can be used which signifies intermediate values, needed to express that a feature value changes during the realization of the segment (see Section 3.2), or when a property is only weakly present (e.g. half nasalized, see Section 3.4.4.2). However, when presence or absence can be predicted on the basis of one or more of the other features, respectively a '0' and '1' are noted. When H & H apply their feature frequency method, redundant positive marked features (1's) are processed as absent.

In our implementation only 0's, ∗'s and 1's are used. The ∗'s and 1's have the same meaning as respectively the ∗'s and +'s of H & H, and a 0 means that either the feature is absent or redundant or both (the −'s, 1's and 0's of H & H).

H & H give five rules, where rule 1 and rule 2 are subdivided in four subrules. Because of the restriction to and the different meaning of 0's, $*$'s and 1's in our implementation, we first have to perform the rules 2a to 2d, and the rules 1a to 1d afterwards. In this way, we get the same effect as when following the operating procedure of H & H exactly. Table 3.17 contains the rules as given by H & H, extended with some more rules, which we will justify below.

The rules 1a to 1e indicate that the features [+front] ... [+polar] predict the feature [+vowel]. For the sake of closing diphthongs we added the rules 1f to 1j, and for the purpose of the centering diphthongs we added the rules 1k to 1o.

The rules 1e, 1j and 1o are added for the sake of sounds marked as [−front −back −round −low +polar]. Sounds with these specifications do not appear in the feature table of H & H because it contains only sounds which appear in the RND. However, in view of the NOS, we extend the system so that all IPA vowels can be processed.

Vowels are always marked as [+sonorant +voiced +continuant +syllabic]. This is indicated in the rules 2a to 2d. We added the condition [−consonant]. Syllabic consonant are also specified as [+vowel], however, since they are also [+consonant], they are excluded by this extra condition, in accordance with what is suggested by the feature table of H & H. For the benefit of the centering diphthongs we added the rules 2e to 2h.

H & H mention that for diphthongs the feature [+long] is redundant. From the feature table it can be concluded that only closing diphthongs are intended to fall under this rule. Centering diphthongs are – perhaps surprisingly – specified as [−diphthong]. The prediction of [+long] for closing diphthongs is reflected in rule 3.

For consonants pronounced in the back the feature [+high] is superfluous: This is reflected in rule 4a. From the feature table follows that also rule 4b applies.

H & H write that for the sonorant laryngeal, the guttural [ɦ] the feature [+continuant] can be predicted. They suggest as rule: [+laryngeal +sonorant] → +continuant. However, in the feature table the [ɦ] is specified as [−sonorant +voiced]. Upon investigation, it appears that as a second condition it was not [+sonorant], but [+voiced] that was meant. In our overview the rule is corrected and given as rule 5.

From the feature table it appears that for vocalized (or syllabic) consonants the feature [+vowel] is redundant. Therefore, we added rule 6.

H & H do not give rules which predict negative marked features since there is no real difference between absent features and redundant features when processing the feature specifications.

1a	+front		→	+vowel
1b	+back		→	+vowel
1c	+round		→	+vowel
1d	+low		→	+vowel
1e	+polar		→	+vowel
1f	∗front	+diphthong	→	+vowel
1g	∗back	+diphthong	→	+vowel
1h	∗round	+diphthong	→	+vowel
1i	∗low	+diphthong	→	+vowel
1j	∗polar	+diphthong	→	+vowel
1k	∗front	−diphthong	→	∗vowel
1l	∗back	−diphthong	→	∗vowel
1m	∗round	−diphthong	→	∗vowel
1n	∗low	−diphthong	→	∗vowel
1o	∗polar	−diphthong	→	∗vowel
2a	+vowel	−consonant	→	+sonorant
2b	+vowel	−consonant	→	+voiced
2c	+vowel	−consonant	→	+continuant
2d	+vowel	−consonant	→	+syllabic
2e	∗vowel	−consonant	→	+sonorant
2f	∗vowel	−consonant	→	+voiced
2g	∗vowel	−consonant	→	+continuant
2h	∗vowel	−consonant	→	+syllabic
3	+diphthong		→	+long
4a	+posterior		→	+high
4b	∗posterior		→	+high
5	+laryngeal	+voiced	→	+continuant
6	+syllabic	+consonant	→	+vowel

Table 3.17: Redundancy rules which predict positive feature specifications as used in the feature system of H & H.

3.6 Comparison of segments

3.6.1 Phones

Calculating sound distances on the basis of the phone representation is trivial. There exist only two distances: 0 (phones are equal) and 1 (phones are different). So the distance between e.g. a [p] and an [a] is 1 unit, the distance between a [p] and a [b] is 1 unit, and also the distance between an [au] and an [a] is 1 unit. In the approach of Kessler (1995) two phones which are basically equal but have different diacritics, are regarded as different phones. So [a] versus [aː] costs 1 unit just as [a] versus [p] costs 1 unit. Our approach only deals with basic symbols. Suprasegmentals and diacritics can only be taken into account by changing the transcription beforehand (see Section 3.4). Using the resulting transcription only the basic symbols are processed. This approach is motivated by the idea that we should take care not to overvalue the influence of suprasegmentals and/or diacritics and retain the relation between a sound with and without one or more additional marks. So in our research an [a] and an [aː] are considered to be equal.

3.6.2 Features

One of the properties of the feature system of H & H is that all features are binary. H & H developed their system for their feature frequency method. In the systems of V & C and A & B also multivalued features are used. A disadvantage of multivalued features is that they may neutralize each other. In this section we describe how a multivalued feature can be changed into a vector of binary features. The neutralizing effect is illustrated on the basis of the three cases where the effect was found. For each of the cases we show how this effect is eliminated when using binary vectors. Finally we describe how the distance between two histograms or between two feature bundles is actually calculated, using the binary vector representation.

3.6.2.1 Vector representation per feature value

Comparing the feature representations in the system of H & H with the systems of V & C and A & B, we see that H & H only use values 0 and 1, represented by + and −. For diphthongs and extra-short sounds also the value 0.5 is used, represented by a ∗. In the two other systems a wider range of values is used. In the three feature systems vowel advancement for instance is defined as follows:

	H & H		V & C	A & B
	front	back	advancement	advancement
front	1	0	2	1
central	0	0	4	2
back	0	1	6	3

Where V & C and A & B use one multivalued feature, H & H use two binary features. In the V & C system the values of the A & B system are weighted two times. The weighting in the H & H system is the same as in the A & B system. In fact H & H give a vector representation for the A & B values. In general the following applies: one feature having n integers with stepsize 1 can always be converted to a vector of $n-1$ binary values. Other possibilities of representing the A & B advancement feature by a binary vector are:

	value 1	value 2
front	0	0
central	1	0
back	1	1

and:

	value 1	value 2	value 3
front	1	0	0
central	1	1	0
back	1	1	1

The first representation is most efficient. However, we prefer the second representation. The idea behind the second is: 1 is represented by one 1, 2 by two 1's, 3 by three 1's, etc. If necessary, the value 0 can also be used and represented as a vector containing only 0's. In that case the following applies: a multivalued feature which may have as its highest value the value n, and which only contains integers, can always be converted to a vector containing n binary features.

As mentioned above, besides the 0 and 1 H & H use also the value 0.5. When processing suprasegmentals and diacritics (see Section 3.4) it appeared that we needed some more fractions. In our research we use the following fractions: 0.125, 0.250, 0.375, 0.500, 0.625, 0.750 and 0.875.

3.6.2.2 Summing feature values

When using the frequency-based corpus frequency method or the frequency per word method per feature the values as specified for the segments in the corpus or in the word are added. When using multivalued features, low and high values may neutralize each other. Assume a hypothetical example where one dialect has one vowel which is fronted and one vowel which is back. Another dialect has two vowels which are central. Using the A & B system, the sum of the values for the feature advancement in the first dialect is $1+3=4$, and in the second dialect $2+2=4$. This suggests erroneously that the two dialects are equal when only considering the feature advancement. When using binary vector representations this will not be the case. We use the vector representation as suggested in the previous section. In this representation the feature advancement is split in three binary features. Now for each dialect the values of the features of the one vector are added to the values of the corresponding features of the other vector:

	v1	v2	v3
front	1	0	0
back	1	1	1
sum	2	1	1

	v1	v2	v3
central	1	1	0
central	1	1	0
sum	2	2	0

Next we calculate the absolute differences between the corresponding sum values of the one dialect and those of the other dialect:

	v1	v2	v3
dialect 1	2	1	1
dialect 2	2	2	0
abs. diff.	0	1	1

The distance between the two dialects is found by taking the sum of the absolute differences: $0 + 1 + 1 = 2$. This outcome shows correctly that the two dialects are different.

3.6.2.3 Representation of diphthongs

As mentioned in Section 3.2, the definition of diphthongs is based on the features corresponding with the start position and with the end position. To be more concrete, the average feature values of the values of the start position and those of the end position are taken. Assume the four degrees of height are defined as 1 (close), 2 (close-mid), 3 (open-mid) and 4 (open). Now we want to find the correct height value for the [ɛi]. The [ɛ] is open-mid (3), the [i] is close (1). The average would be $(1 + 3)/2 = 2$, i.e. close-mid. However, the [e] is close-mid is well. So we get a neutralizing effect when taking the average feature values, resulting in a specification which suggests that the [ɛi] has the same height as the [e]. Since the [e] is a stable sound, and the [ɛi] is a sound with a changing height, this outcome is undesirable. We found the solution by representing the multivalued feature height as a binary vector. Analogous to the example in Section 3.6.2.1 the different degrees of height are represented as follows:

	v1	v2	v3	v4
close	1	0	0	0
close-mid	1	1	0	0
open-mid	1	1	1	0
open	1	1	1	1

Using these representations, the values of the corresponding features of the two vectors are averaged:

	v1	v2	v3	v4	
[i]	1	0	0	0	×50%
[ɛ]	1	1	1	0	×50%

gives:

	v1	v2	v3	v4
	0.5	0	0	0
	0.5	0.5	0.5	0
[ɛi]	1	0.5	0.5	0

Next we calculate the absolute differences between the corresponding vector values of the [ɛi] and the [e]:

	v1	v2	v3	v4
[e]	1	1	0	0
[ɛi]	1	0.5	0.5	0
abs. diff.	0	0.5	0.5	0

The sum of the differences is $0 + 0.5 + 0.5 + 0 = 1$, which shows that the stable [e] and the changing [ɛi] are different.

3.6.2.4 Representation of the place of articulation

Using vector representations for multivalued features is also useful when finding the representation of the place of articulation of a sound with a secondary place of articulation. In our research the primary place of articulation is weighted for 75% and the secondary place of articulation for 25% (see Section 3.4). Assume we have to process a velarized t ([tˠ]). The place of articulation of the [t] is alveolar. Assume alveolar is represented by 1, postalveolar by 2, retroflex by 3, palatal by 4 and velar by 5. Using a single value, the new place of articulation would be $75\% \times 1 + 25\% \times 5 = 2$, which is postalveolar. However, a velarized t is not postalveolar at all. Here we are faced again with the difficulty that values of multivalued features neutralize each other. With the use of vector representations this problem will be solved. Analogous to the example in Section 3.6.2.1 the different places of articulation are represented as follows:

	v1	v2	v3	v4	v5
alveolar	1	0	0	0	0
postalveolar	1	1	0	0	0
retroflex	1	1	1	0	0
palatal	1	1	1	1	0
velar	1	1	1	1	1

Using this representations, we should weight the primary place of articulation 75% (alveolar) and the secondary place of articulation (velar) 25%. When adding the weighted vector values of the primary place of articulation to the corresponding weighted vector values of the secondary place of articulation, we get the place of articulation of the velarized t:

	v1	v2	v3	v4	v5	
alveolar	1	0	0	0	0	×75%
velar	1	1	1	1	1	×25%

gives:

	v1	v2	v3	v4	v5
	0.75	0	0	0	0
	0.25	0.25	0.25	0.25	0.25
sum	1	0.25	0.25	0.25	0.25

Next we calculate the absolute differences between the corresponding vector values of the velarized t and postalveolar:

	v1	v2	v3	v4	v5
velarized t	1	0.25	0.25	0.25	0.25
postalveolar	1	1	0	0	0
abs. diff.	0	0.75	0.25	0.25	0.25

The sum of the differences is $0 + 0.75 + 0.25 + 0.25 + 0.25 = 1.5$, which shows that the place of articulation of a velarized t is distinguished from postalveolar.

3.6.2.5 Comparison of segments

The comparison of feature histograms and feature bundles is basically the same. Therefore, the same metric for the comparison of feature histograms is used for the comparison of feature bundles as well. There are several metrics for finding the distance between two feature histograms or feature bundles (Jain and Dubes, 1988; Hoppenbrouwers and Hoppenbrouwers, 1988). We restricted ourselves to the most common ones: Manhattan (or taxicab, or city block) distance, Euclidean distance and Pearson's correlation coefficient.

Assume we compare a histogram or bundle X with a histogram or bundle Y where n is the number of features. The Manhattan distance (Jain and Dubes, 1988) is simply the sum of all feature value differences for each of the n features:

$$\delta(X,Y) = \sum_{i=1}^{n} |X_i - Y_i| \tag{3.1}$$

The Euclidean distance is the square root of the sum of squared differences in feature values:

$$\delta(X,Y) = \sqrt{\sum_{i=1}^{n} (X_i - Y_i)^2} \tag{3.2}$$

The Pearson correlation coefficient (Hogg and Ledolter, 1992) is calculated as follows:[8]

$$r(X,Y) = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \overline{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \overline{X})^2}} \tag{3.3}$$

When comparing two ranges, if all values are equal in one or both histograms, the correlation coefficient between both ranges is not defined since a division by zero occurs. This never happens in corpus-based histograms, but it is possible that all values in a word-based histogram or in a single feature bundle are the same. Normally the correlation coefficient ranges from -1 (inverse ranges) to $+1$ (parallel ranges). Therefore, when both feature bundles are constant, we set the correlation coefficient to 1. When only one range is constant, we set the correlation coefficient to 0.

In fact, Pearson's correlation coefficient is a similarity measure. As such Hoppenbrouwers and Hoppenbrouwers (2001) used this metric for the comparison of feature histograms. The minimum value is -1 (minimal similarity) and the maximum value is $+1$ (maximal similarity). In view of the use of cluster analysis (see 6.1) we only want to use distance metrics. We used the Pearson's correlations coefficient by calculating $1 - r$. In that case the minimum value is 0 (maximal similarity) and the maximum value is 2 (minimal similarity).

We are aware of the fact that the use of the Pearson's correlation coefficient is more correct for the comparison of histograms than for the comparison of feature bundles. Histograms which are parallel to each other show that in the corresponding varieties the different features are positively marked in the same proportions. However, when comparing feature bundles this approach may give the wrong results when ranges are parallel and the one range is consistently higher than the other range. The Pearson's correlation coefficient will give the impression that they are equal. We include measurements of correlation coefficients among feature bundles for the sake of completeness, even though we do not expect it to function well.

---

[8]Before using this formula, $\overline{X}$ and $\overline{Y}$ should be calculated. This means that for calculating the correlation coefficient at least two passes are needed. In our research we used another formula, which allowed more efficient processing and avoids some of the rounding errors that are made with the earlier formula (Hogg and Ledolter, 1992):

$$r(X,Y) = \frac{\sum_{i=1}^{n} X_iY_i - \frac{\sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i}{n}}{\sqrt{(\sum_{i=1}^{n} X_i^2 - \frac{(\sum_{i=1}^{n} X_i)^2}{n})(\sum_{i=1}^{n} Y_i^2 - \frac{(\sum_{i=1}^{n} Y_i)^2}{n})}}$$

3.7 Linear and logarithmic distances

Using the feature bundles the distance between two segments can be calculated. The simplest way in which this can be done is taking the sum of the absolute differences of each pair of corresponding feature values. Other metrics are described in Section 3.6.2. Characteristic for the original A & B system is that when a distance exceeds a certain ceiling, that distance is set to the value of a ceiling. However, the question arises as to what value the ceiling should be set to. In Heeringa and Braun (2003) an adjusted version of the Almeida & Braun system is proposed where the logarithm of the feature bundle distances is taken instead of using a ceiling. The effect of taking the logarithm is that small distances are weighted relatively more heavily than large distances. This may be in accordance with our perception, where small differences in pronunciation play a relatively strong role in comparison with larger differences. We experimented with both linear and logarithm feature bundle distances for all three feature systems: the H & H, the V & C and the A & B system.

Because the distance between identical sounds is 0, and the logarithm of 0 is not defined, we first increase the distance with 1 and next calculate the logarithm of the distance. In this way, the distance between equal sounds still remains 0 since the logarithm of 1 is equal to 0. In general we calculate $ln(distance + 1)$.

In Figure 3.2 the effect of taking the logarithm of the IPA vowel distances as found with the A & B system is shown. For each of the 28 IPA vowels the distance with respect to silence is calculated. Next the distances are sorted from short to long. In both cases, linear and logarithmic, [ə] is most like silence and [i], [y], [ɯ], [u], [a], [ɶ], [ɑ], [ɒ] are all most unlike silence. The graph shows the sorted distances. The points corresponding with distances are connected by lines to get a clearer picture. By taking the logarithm, greater distances are decreased to a greater degree than short distances.

In Figure 3.3 the effect of taking the logarithm of the IPA consonant distances as found with the A & B system is shown. For each of the 59 IPA consonants the distance with respect to silence is calculated. Next the distances are sorted from short to long. In both linear and logarithmic distance, [ʔ] is most like silence and [w] is most unlike silence. The graph shows sorted distances in the same way as was shown for the vowels. The points corresponding with distances are connected by lines. Of course the same effect as for the vowels is seen here: greater distances are decreased to a greater degree than shorter distances when taking the logarithm.

We only apply the logarithm to segment distances as used in the Levenshtein distance (see Section 5.1), not to histogram distances as calculated in the corpus frequency method or the frequency per word method. In the corpus feature frequency method the distance between two histograms corresponds to a dialect distance, and in the feature frequency per word method the distance between two histograms corresponds to a word distance (see Section 2.3.2 and Section 2.3.3).

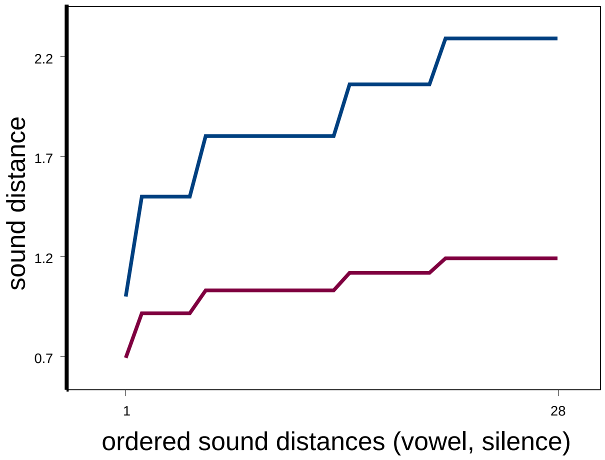

Figure 3.2: Linear (upper) and logarithmic (lower) A & B distances of 28 IPA vowels with respect to silence. Distances are calculated as the sum of the differences between corresponding features. The graph shows the distances sorted from low (left) to high (right). Greater distances are reduced more than smaller ones by using the logarithm.

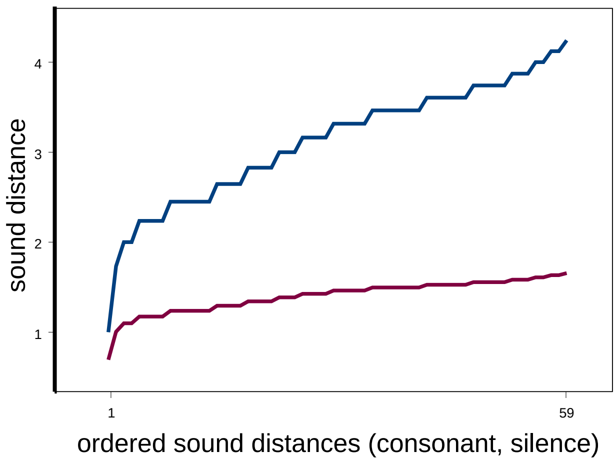

Figure 3.3: Linear (upper) and logarithmic (lower) A & B distances of 59 IPA consonants with respect to silence. Distances are calculated as the sum of the differences between corresponding features. The graph shows the distances sorted from low (left) to high (right). Greater distances are reduced more than smaller ones by using the logarithm.

The result of taking logarithmic distances is that great dialect distances (in the corpus feature frequency method) or great word distances (in the feature frequency per word method) will stay relatively small. For dialect distances and word distances we judge this to be undesirable.

Special attention should be paid when the [i], [j], [u] and [w] are compared to each other when using the feature systems of V & C and A & B. In Section 3.1.3.1 we described that these sounds are defined as both vowel and consonant. When a doubly defined segment is compared to another doubly defined segment, the distance is equal to the average of the vowel distance and the consonant distance. When calculating the logarithmic distance, we calculate this as:

$$\frac{ln(vowel\ distance + 1) + ln(consonant\ distance + 1)}{2}$$

3.8 Correlation between systems

In this section we compare the different feature systems of H & H (see Section 3.1.2), V & C (see Section 3.1.3) and A & B (see Section 3.1.4). For each of the systems all vowels and consonants are specified. We ask the question to what degree the various systems for calculating segment distances correlate in the distances they assign, and whether there remain interesting differences as well. The distances between vowels and consonants are calculated using the three metrics which are proposed in Section 3.6.2: the Manhattan distance (M.), the Euclidean distance (E.) and the Pearson correlation coefficient (P.). For the H & H system redundancy is removed from the feature bundles (see Section 3.5). For all systems the linear distances are used (see Section 3.7).

In Section 3.8.1 we give an overview of all matrices among which the correlation coefficients are calculated. Section 3.8.2 explains how to determine if a correlation coefficient is significant. In this section we also explain how to determine whether two correlation coefficients are significantly different. In Section 3.8.3 we examine the influence of the different feature systems while in Section 3.8.4 the influence of different feature bundle metrics is investigated.

In Section 7.4.3 the influence of different feature systems and different feature bundle metrics will be further examined.

3.8.1 Matrices

When correlating results of one particular segment representation using one particular feature bundle metric with results of another segment representation using a feature bundle metric, this is done on the basis of distances which are arranged in a matrix. The matrices may have four different sizes.

In the RND data 18 vowels are used. Although the vowel [œ] does not appear in RND transcriptions, it is included because it is used for the definition of diphthongs (see Table 3.14). Calculating the distances between the vowels results in $(18 \times 17)/2 = 153$ distances. In the RND 27 consonants are used. When calculating the distances between the consonants we get $(27 \times 26)/2 = 351$ distances.

In the NOS data the modern IPA system is used. In the IPA system 28 vowels are given. If we calculate the distances between the vowels, we get a distance matrix of $(28 \times 27)/2 = 378$ distances. In the IPA system 58 pulmonic consonants are given. We added the [w] which is ordered under 'Other Symbols' in the IPA system, so we get 59 consonants. When calculating the distances between the consonants, a distance matrix of $(59 \times 58)/2 = 1711$ distances is obtained.

3.8.2 Significance

For finding the significance of a correlation coefficient we used the Mantel test. In classical tests the assumption is made that the objects which are correlated are independent. However, values in distance matrices are usually correlated in some way, and not independent (Bonnet and Van de Peer, 2002). A widely used method to account for distance correlations is the Mantel test (Mantel, 1967). Mantel developed an asymptotic test, in which the null hypothesis is that distances in the one matrix are independent of the corresponding distances in the other matrix. The significance of the statistic can also be evaluated by randomly reallocating the order of elements in one of the matrices (Bonnet and Van de Peer, 2002).

The program we used is also based on a series of random permutations. Assume we have two matrices D_1 and D_2. We would like to know whether $r(D_1, D_2)$ is significant. To determine the significance a number of iterations is performed. In each iteration the order of the elements of D_1 is changed by swapping each element with another element where the other element is randomly chosen. In fact it does not matter whether D_1, D_2 or both are randomly permuted. Next the following condition is tested:

$$r(P_1, D_2) > r(D_1, D_2)$$

If this condition is true, a counter is incremented by 1. After the iterations are finished, the counter is divided by the number of the iterations. The outcome gives the chance that randomly permuted matrices correlate more strongly than the two unchanged matrices. The number of iterations determine the overall precision of the test. Since we use $\alpha = 0.05$, the number of repetitions should be equal to about 1000 (Manly, 1997).

Besides finding the significance of a correlation coefficient we would like to know whether one correlation coefficient is significantly higher than another. Assume we have four matrices D_1, D_2, D_3 and D_4. We want to know whether $r(D_1, D_2)$ is significantly higher than $r(D_3, D_4)$. For this purpose, a number of iterations is performed again. In each iteration D_1 and D_3 are randomly permuted. We call the permuted matrices respectively P_1 and P_3. Just as in the procedure described above, a random permutation is generated by swapping each element with another element which is randomly chosen. Next the following condition is tested:

$$r(P_1, D_2) - r(P_3, D_4) \geq r(D_1, D_2) - r(D_3, D_4)$$

If the condition is true, a counter is incremented by one. After all iterations are finished, the counter is divided by the number of the iterations. This gives the chance of getting a difference which is equal to or greater than the given difference when using randomly permuted matrices. Just as in the previous procedure, we perform 1000 iterations and a significance level of $\alpha = 0.05$.

Examples of related applications of the Mantel test are found in Barbujani et al. (1994), Weng and Sokal (1995) and Manni (2001). Barbujani et al. (1994) investigated the relation between genetics and linguistics in the Caucasus. Genetic, geographic and linguistic distances were correlated. Weng and Sokal (1995) carried out a lexicostatistics study. In this study a series of tests was undertaken to relate lexicostatistical dissimilarities among 48 Indo-European languages to distances representing various causal hypotheses. The putative causal distance matrices include geographic distances, distances representing the origin of agriculture, and distances representing hypotheses concerning the origin and spread of Indo-European languages in Europe. Manni (2001) compared genetic and linguistic distances for the Italian province of Ferrara and for the Netherlands.

3.8.3 Feature representations

In Tables 3.18 and 3.19 the different feature representations can be compared. In the tables correlation coefficients are given, based on pairs of matrices where the distances in each matrix are based on different feature systems and on the same metric for the comparison of feature bundles. The columns divide results in Manhattan, Euclidean and Pearson metrics. Results are given for vowels and consonants for each metric. All correlations are significant for $\alpha = 0.05$.

For both the RND and the IPA it appears that all correlations between the V & C system and the A & B system are stronger – although not significantly stronger – than the corresponding correlations between any other pair of systems. Looking at the vowel features, height is defined in a similar way in the V & C system and the A & B system, while in the H & H system some relationship between most low and most high sounds is defined. For the consonants it holds that in both the V & C system and the A & B system the place of articulation is explicitly defined. This is not the case in the H & H system. In turn most correlations between the H & H system and the V & C system are stronger – but not significantly stronger – than the corresponding correlations between

	M.		E.		P.	
	vow.	cons.	vow.	cons.	vow.	cons.
H & H vs. V & C	0.77	0.68	0.77	0.71	0.59	0.68
H & H vs. A & B	0.70	0.58	0.72	0.60	0.51	0.65
V & C vs. A & B	0.79	0.80	0.80	0.81	0.67	0.77

Table 3.18: Correlation coefficients between RND segment distances obtained on the basis of different feature systems and as calculated using Manhattan (M.), Euclidean (E.) and Pearson (P.) procedures. Results are given per metric for vowels (vow.) and consonants (cons.).

	M.		E.		P.	
	vow.	cons.	vow.	cons.	vow.	cons.
H & H vs. V & C	0.72	0.61	0.71	0.62	0.42	0.63
H & H vs. A & B	0.64	0.57	0.65	0.61	0.42	0.64
V & C vs. A & B	0.80	0.76	0.79	0.76	0.71	0.74

Table 3.19: Correlation coefficients between IPA segment distances obtained on the basis of different feature systems and as calculated using Manhattan (M.), Euclidean (E.) and Pearson (P.) procedures. Results are given per metric for vowels (vow.) and consonants (cons.).

the H & H system and the A & B system. In the A & B system, manner of articulation is defined as a scale. This is not the case in the two other systems.

The correlation coefficients vary from at least 0.42 to at most 0.81. These rather low correlation coefficients show that it remains interesting to take all three feature representations into account in further research, even though the correlation coefficients are significant.

3.8.4 Feature bundle metrics

In Tables 3.20 and 3.21 the different metrics can be compared. In the tables correlation coefficients are given, based on pairs of matrices where the distances in each matrix are based on different metrics for feature bundle comparison and on the same feature system. The columns divide results in the H & H, V & C and A & B system. All correlations are significant for $\alpha = 0.05$.

For both the RND and the IPA it appears that all correlations between the Manhattan metric and the Euclidean metric are stronger than the corresponding correlations between any other pair of metrics. Examining the IPA results, the correlations between the Manhattan metric and the Euclidean metric are also significantly stronger when using the feature systems of H & H (vowels and con-

	H & H		V & C		A & B	
	vow.	cons.	vow.	cons.	vow.	cons.
M. vs. E.	0.97	0.99	0.98	0.99	0.99	0.98
M. vs. P.	0.89	0.90	0.90	0.94	0.94	0.95
E. vs. P.	0.87	0.90	0.90	0.94	0.94	0.93

Table 3.20: Correlation coefficients between RND segment distances obtained on the basis of different metrics. Results are given per feature system for vowels (vow.) and consonants (cons.).

	H & H		V & C		A & B	
	vow.	cons.	vow.	cons.	vow.	cons.
M. vs. E.	0.97	0.97	0.98	0.98	0.99	0.99
M. vs. P.	0.80	0.90	0.90	0.94	0.95	0.94
E. vs. P.	0.80	0.89	0.89	0.93	0.95	0.93

Table 3.21: Correlation coefficients between IPA segment distances obtained on the basis of different metrics. Results are given per feature system for vowels (vow.) and consonants (cons.).

sonants) and A & B (consonants). We observe that some correlations between the Manhattan metric and the Pearson correlation are higher than the corresponding correlations between the Euclidean metric and the Pearson correlation. However, these correlations are not significantly higher.

When looking at the correlation coefficients between the Manhattan distances and the Euclidean distances, we see that these vary from at least 0.97 to at most 0.99. These are all extremely high values, indicating that it is not necessary to regard both metrics in further research. The correlation coefficient between the Manhattan metric and the Pearson metric vary from 0.80 to 0.95, and between the Euclidean metric and the Pearson metric from 0.80 and 0.95. This indicates that the Pearson correlation coefficient metric is rather different from both other metrics.

3.9 Conclusions

In this chapter we proposed different representations of speech segments in order to find the relations among them. The roughest representation is the *phone* representation. In this representation speech segments are simply equal or not equal. There are no gradations. The more refined representation is the *feature* representation. Using this representation, finer segment distances are obtained. We investigated three feature systems, and three metrics for finding distances between feature histograms (used in frequency-based methods) and feature bundles (used in Levenshtein distance).

We found that the V & C system and the A & B system correlate for most metrics more strongly – although not significantly so – than the corresponding correlations between any other pair of systems. For both systems the vowel feature which defines the height is defined in a similar way and the place of articulation is explicitely defined for consonants. The rather low correlations between all three systems show that it remains interesting to take all three feature representations into account in further research, although the correlation coefficients are significant.

It appeared that the correlations between the Manhattan metric and the Euclidean metric are for all feature systems stronger than the corresponding correlations between any other pair of metrics. Some of them were also significantly stronger. The strong correlation indicates that it is actually not necessary to consider both metrics in later work. The Pearson correlation coefficient appeared to be rather different from the two other metrics.